DOCUMENT RESUME

ED 091 416 TM 003 630

AUTHOR Rosenbach, John H.
TITLE Achievement vs. Learning: Needed Distinctions?
PUB DATE [73]
NOTE 8p.; Paper presented at the Annual Meeting of the Northeastern Educational Research Association (Ellenville, N.Y., November, 1973)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS *Achievement; Achievement Tests; Aptitude Tests; *Intelligence; *Learning; Performance Criteria; *Testing Problems

ABSTRACT

Since at least 1897, educational researchers have reported, with great frequency, "no significant differences" with respect to studies investigating the effects on educational outcomes of various treatments, suggesting that many educational variables are relatively impotent. A possible reason for at least many of the no-difference findings may lie in the nature of the criteria and especially in the use of standardized achievement tests as criterion measures. Oftentimes implicit in the use of such tests are the assumptions that learning and achievement are equivalent concepts and that achievement is modifiable through instruction. Given that educational research findings are now being used as a basis for formulating state and national policies, it becomes increasingly important that greater attention be given to exploring why so many studies report no-differences. Therefore it is suggested that (1) the distinctions between achievement and intelligence, if any, be clarified, (2) efforts be made to develop instruments which are capable of detecting unique contributions to the school (primarily instruction) to changes in the students, and (3) attempts be made to understand both the similarities and differences between the concepts of learning and achievement. (Author/RC)

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM 003 630 ED 091416

Achievement vs. Learning: Needed distinctions?*

John H. Rosenbach

SUNY at Albany

Concern over the adequacy of criteria used to judge human performance is hardly new. Researchers and others involved in evaluation have, for years, bemoaned the problems of finding reliable, valid, and useful criterion measures. Just such concerns, in fact, were at least in part responsible for the cooperative efforts of APA, AERA, and NCME which led to the publication of the various "Technical Recommendations" (1954, 1955) and "Standards" (1966, 1973) for development and use of educational and psychological tests. The earlier documents undoubtedly led to more technically sound instruments and to more sophisticated test interpretation. But developments with respect to improved criteria seem to have lagged behind as illustrated by the relatively recent furor which continues to build around the issue of job-related criteria. It is quite clear that once we move beyond predicting educational achievement, our available instruments prove quite weak; when test procedures are applied to job selection, not only does predictive accuracy drop but serious questions of relevancy and discrimination arise, as well.

It is surprising that similar cries have not been more forcefully and systematically raised with respect to criteria used to evaluate educational treatments, given the long history

---

*Presented as part of a symposium, "Criteria and the Evaluation of Educational Programs." Annual meeting of the Northeastern Educational Research Association, Ellenville, N.Y., November 2, 1973.

of research suggesting the relative impotency of so many educational variables [e.g., Astin (1968), Coleman (1966), Cook (1951), Rice (1897), Stephens (1967)]. To be sure, people, especially teachers and administrators, complain that the tests do not measure those "important things we are really teaching" or simply dismiss the entire issue by proclaiming "pencil and paper tests aren't any good, anyway." But there has been little concerted effort to address the question as to why educational research continues to yield the same results [Stephens (1967) is a notable exception] and the related question, might there be something about the nature of the criteria themselves (usually standardized achievement tests) that tends to produce these findings?

About three years ago or so I had just about reached the point of accepting the no-difference findings at face value; apparently, variations in educational treatments did not bring about differential educational outcomes. But, be it intuition, belief, tradition, or what, there was something about this conclusion that didn't quite ring true. But where was the ringer? In the treatments? The system? Clumsy and imprecise instruments? Or, perhaps, in what we were looking at as criteria?

One clue to this puzzle surfaced when I began to reflect on the purposes of schooling. Up to this point I had tacitly assumed, as Cook (1951) so succinctly put it, "The central problem of all educational endeavors is learning [p 3]." But when one examines closely what schooling is all about and what teachers and even parents show most concern for, the centrality of learning becomes less evident. What does emerge from such an analysis is the primacy of achievement, i.e., what a person can do at a given

time; there is little regard shown for how a person has changed over a period of time. Several questions, some philosophical, follow from this observation: What should be the central concern of schooling? Should greater recognition be given to change, thus placing more emphasis on learning? Should the reward (and punishment) system be based on status (achievement) or improvement? What is the relationship of achievement and learning? To what extent are achievement tests measures of learning? To what degree do achievement tests reflect the contributions (or even potential contributions) of school processes, as contrasted with personal variables and/or pre- or out-of-school influences, to the trait or characteristic being measured?

The questions of a "should" nature do not fall within the province of behavioral science. But those remaining are clearly within the domain of educational research. In fact, they arise, in part, because of the work of educational researchers over the past fifty years. What is shocking, however, is, given the widespread use of achievement tests to evaluate individual pupils, programs, schools, and in some cases, even teachers, that these questions still remain unanswered. What is especially troubling is that without this knowledge we have no way of intelligently assigning, or even estimating, responsibility for educational progress. This is not to say, of course, that this has prevented, or will prevent, us from engaging in just such actions.

I became painfully aware of the ramifications of this issue while serving as a consultant to the New York State Education Department. I spent several weeks observing elementary classes in New York City that were partially funded through State

and Federal sources. Somewhat to my surprise (indicative, I'm sure of Upstate mentality) I came away extremely impressed with what I had seen. In school after school I observed dedicated, professional teachers, interesting and well-conceived instructional programs, well-trained paraprofessionals, and a wealth of instructional and technological resources (and, perhaps most surprisingly, a lot of happy kids). Yet, when I queried district administrators about the success of these programs, the most optimistic response was the poignant remark, "Well, at least our test scores are going down at a rate slower than that of the rest of the City!"[1] To those of us who have become conditioned to the notion of input-output relationships, this statement is hardly a surprise; but to those legislators and administrators who are responsible for funding, it is a devastating shock, for they had been led to believe that massive spending would indeed produce discernible differences. How do we explain these results to the profession and to the public?

At this point, I was reminded of the words of Truman Kelley written in 1927. Kelley coined the "jangle fallacy," or "...the use of two separate words or expressions covering in fact the same basic situation, but sounding different, as though they were in truth different [Kelley, 1927, p 65]." What he was referring to were the concepts of intelligence and achievement (as measured by standardized tests). He raised the questions: "How much of achievement is intelligence?" and "How much of intelligence is

---

[1]Recently, the New York Times has reported that the 1973 M.A.T. scores have, for the first time in years, shown gains.

achievement?" And then answered by stating "...no less than 90 per cent of the one is the same in its nature as the other [p 62]." He then went on to argue that to classify people "...upon the basis of their difference in these two traits is a sheer absurdity [p 65]." Forty-six years later we continue to act absurdly! For we have come to admit the inevitably of input determining output by failing to recognize that the input and output variables are the same thing; we have simply succumbed to the "jangle fallacy." If we were asked, as educators, to modify intelligence (even simply defined as scores on an IQ test), we would respond humbly; we know, from years of research and on the basis of how IQ tests are constructed, that this construct is almost impossible to change. But give it a different name, "achievement," and we are eager to demonstrate what we can accomplish (but don't!). We realize that "vocabulary," as the single best indicator of general intelligence, is virtually impossible to improve, but we are happy to work toward enhancing "word recognition" on a reading achievement test; we know that "comprehension," as a major factor related to "g" is virtually immutable, but we design programs in "problem solving" or "reading comprehension." And we exhibit surprise and consternation when the programs don't work!

Even if we had never read Kelley (1927), the more contemporary works of Bloom (1964) or of Coffman (1969), who describe the tremendous consistency of achievement test scores, should have provided clues. But perhaps the most absurd of all, is our acceptance of the Coleman study (1966), in which aptitude tests were admittedly used as criterion measures! And what's so profound about the finding that school variables (e.g.,

expenditures, books in library, teacher quality, and so forth) have little bearing on aptitudinal differences? Yet, such results have led to massive changes in educational funding.

Our negligence as professional educational researchers has not only academic, theoretical implications, but fundamental societal implications, as well. And it is time that we recognize these responsibilities.

Where do we go from here? First, by rejecting the self-serving notion that "...it is only within recent years that the full power of measurement to modify and improve instructional procedures has been realized... [Cook, 1951, p 6]." Secondly, by recognizing the need to develop instruments to detect the unique contributions of the school (or any other variable, for that matter) to educational progress. Very likely, our direction must be toward more criterion-referenced like approaches. Robert Glaser, in 1963, has spelled out some of what is required:

> Such measures [norm-referenced] need provide little or no information about the degree of proficiency exhibited by the tested behaviors in terms of what the individual can do. They tell that one student is more or less proficient than another, but do not tell how proficient either of them is with respect to the subject matter tasks involved...achievement tests used...to provide information about differences in treatments need to be constructed so as to maximize the discriminations made between groups treated differently and to minimize the differences between the individuals in any one group.... The content of the test used to differentiate treatments should be maximally sensitive to the performance changes anticipated from the instructional treatments. (p 520)

Finally, we must move toward a clearer understanding of achievement (actually multiple concepts related to all achievement domains) as a hypothetical construct, and especially to its relationship to learning. One of the most puzzling facets of this

problem is that while, undeniably, learning is a necessary aspect of achievement, the roles of other variables, especially aptitudes and abilities, are obscure. For example, what is the nature of a difference score (aside from error) which is obtained by subtracting a pre-test score from a post-test score. In what ways do aptitude, ability, and treatment variables interact to produce this difference. To assume that this difference is learning is as simple-minded and unjustified as assuming that achievement per se is the product solely of learning.

What is required are advances in our conceptualizations of the concept of achievement, learning, and development, and how they interrelate. It seems naive to hope that significant improvements in educational processes (be they instruction, technology, measurement, or evaluation) will be made without first developing sounder bases of knowledge, both theoretical and empirical.

References

Astin, A. Undergraduate achievement and institutional excellence. Science, 16 August, 1968, 661-669.

Bloom, B. S. Stability and Change in Human Characteristics. New York: Wiley & Sons, 1964.

Carroll, J. B. A model of school learning. Teachers College Record, 1963, 64, 723-733.

Coffman, W. E. Concepts of achievement and proficiency, in Toward a Theory of Achievement Measurement, 1969 Invitational Conference on Testing Problems. Princeton: E.T.S.

Coleman, J. S. et al. Equality of Educational Opportunity. U. S. Department of Health, Education and Welfare, 1966.

Cook, W. The functions of measurement in the facilitation of learning, in E. Lindquist (Ed.) Educational Measurement.

Glaser, R. Instructional technology and the measurement of learning outcomes: some questions. American Psychologist, 1963, 18, 519-521.

Kelley, T. L. Interpretation of Educational Measurements. New York: World Book, 1927.

Rice, J. M. The futility of the spelling frind. The Forum, 1897, 23, 163-172.

"Standards for Development and Use of Educational and Psychological Tests" (Third draft). APA Monitor, February 1973.

Stephens, J. M. The Process of Schooling. New York: Holt, Rinehart & Winston, 1967.